

## ABSTRACT

**Objectives.** This study sought to determine primary sources of data for electronic birth certificates.

**Methods.** A survey was administered from 1997 through 1998 to maternity facilities in New Jersey requesting information about what primary information sources were used for 53 electronic birth certificate variables. Potential information sources included the facilities' maternal and infant medical records, the prenatal record, and a parent-completed birth certificate worksheet.

**Results.** Among the 66 maternity facilities responding, there was significant variation in the choice of primary data sources for the electronic birth certificate variables examined.

**Conclusions.** The variability of primary sources for electronic birth certificate data acquisition represents a potential cause of systematic error in reported vital statistics information. (*Am J Public Health.* 2001;91:814–816)

# New Jersey's Electronic Birth Certificate Program: Variations in Data Sources

John C. Smulian, MD, MPH, Cande V. Ananth, PhD, MPH,  
Maryellen L. Hanley, MD, MPH, Robert A. Knuppel, MD, MPH,  
Judy Donlen, DNSc, RN, and Lakota Kruse, MD, MPH

To facilitate the transfer of vital statistics information in a timely and complete fashion, the New Jersey Department of Health and Senior Services, in cooperation with a variety of organizations in New Jersey, has developed a customized direct-entry extended electronic birth certificate produced by Genesis Systems, Inc (Lewistown, Pa). During the development phase, a data dictionary standardizing definitions for each data field was provided to maternity facilities. Specific procedures for collection and entry of information, including determination of sources of the data used for entry into the electronic birth certificate fields, were left to each individual maternity facility.

Because the sources of data for the electronic birth certificate have the potential to affect state and national vital statistics, we designed this study to evaluate where and how each maternity facility in New Jersey obtained specific input data. We also assessed whether any of the facilities performed routine data validation.

## Methods

From 251 electronic birth certificate data fields, we identified 53 specific fields that we believed to be of greatest interest for study. For each of these priority fields, the survey asked each facility to identify the primary source(s) of the information. Fields potentially having

more than 1 response, such as maternal risk factors, obstetric procedures, maternal complications, and abnormal infant conditions and diagnoses, were treated as a single field.

Four possible choices were provided for primary data sources: (1) hospital maternal medical record, (2) infant medical record, (3) physician office prenatal record, and (4) a worksheet completed by the parent(s) before discharge from the facility after delivery. No

John C. Smulian, Maryellen L. Hanley, and Robert A. Knuppel are with the Center for Perinatal Health Initiatives and the Division of Maternal-Fetal Medicine, University of Medicine and Dentistry of New Jersey–Robert Wood Johnson Medical School/Saint Peter's University Hospital, New Brunswick. Cande V. Ananth is with the Center for Perinatal Health Initiatives and the Division of Epidemiology and Biostatistics, University of Medicine and Dentistry of New Jersey–Robert Wood Johnson Medical School/Saint Peter's University Hospital. Judy Donlen is with the Southern New Jersey Maternal and Child Health Consortium, Pennsauken. Lakota Kruse is with the Maternal and Child Health Epidemiology Program, New Jersey Department of Health and Senior Services, Trenton.

Requests for reprints should be sent to John C. Smulian, MD, MPH, UMDNJ–RWJMS/Saint Peter's University Hospital, 254 Easton Ave, MOB-4th Floor, New Brunswick, NJ 08903-0591 (e-mail: smuliac@umdnj.edu).

This brief was accepted July 20, 2000.

**Note.** Opinions, views, and conclusions expressed in this report are those of the authors and not necessarily those of the Robert Wood Johnson Foundation.

judgments were made as to whether there was an optimum data source.

The survey was administered to electronic birth certificate managers from each of the 71 maternity facilities in New Jersey through 7 regional maternal and child consortia. Managers were responsible for returning completed surveys but may have delegated the task of filling in the information to others. Each facility was asked whether it performed any validation of actual electronic birth certificate data to ensure accuracy. Information regarding the person(s) performing data entry was collected.

Results of the analysis were expressed as the percentage of facilities using a specific data source (maternal medical record, prenatal record, infant medical record, or worksheet) as the primary source of information for each electronic birth certificate field. The presence or absence of a data validation process was evaluated on the basis of the number of annual deliveries (unpaired *t* test) and the level of acuity of obstetric services, level 1 representing basic care, level 2 representing intermediate care, level 3 representing intensive care, and level 4 representing the regional perinatal centers (contingency tables).

## Results

Of the 71 maternity facilities in New Jersey, 66 (93%) returned the survey. Individuals completing the survey included electronic birth certificate personnel in 32 of these facilities (48.5%), information specialists in 7 (10.6%), unit nurse managers in 14 (21.2%), unit secretaries in 11 (16.7%), and unspecified individuals in 2 (3%). Table 1 shows the primary data sources for maternal and infant identifier information, demographics, and payer status. The primary data sources for medical history and outcome information are reported in Table 2.

A total of 28 of 62 facilities (45.2%) reported some type of validation process. Level 1 (6 of 15; 40%) and level 2 (14 of 33; 42.4%) facilities were less likely to perform validations than were level 3 facilities (3 of 5; 60%) and regional perinatal centers (5 of 9; 55.6%), but the difference was not statistically significant (*P* = .31). Facilities performing validation had an average of 1553 (SD=998) annual deliveries, as compared with 1334 (SD=1157) among those without validation procedures (*P* = .45). The primary data-entry personnel were unit secretaries in 20 (30.3%) facilities, dedicated electronic birth certificate personnel in 26 facilities (39.4%), nursing staff in 7 facilities (10.6%), medical records personnel in 1 facility (1.5%), physicians in 1 facility (1.5%), and other miscellaneous personnel in 11 facilities (16.7%).

**TABLE 1—Electronic Birth Certificate Primary Data Sources for Maternal and Infant Identifiers, Demographic Characteristics, and Payer Information: New Jersey Maternity Facilities, 1997–1998 (n=66)**

	MMR, %	IMR, %	PNR, %	WS, %	Missing, No.
<b>Maternal identifier</b>					
Last name	63	...	13	24	5
Maiden name	27	...	5	68	5
Medical record number	86	...	3	11	4
Birth date	56	...	16	28	6
Social Security number	39	...	3	58	5
<b>Infant identifier</b>					
Last name	7	10	...	83	8
Date of birth	48	33	...	19	9
Time of birth	50	33	...	17	9
Medical record number	6	83	...	11	4
<b>Demographics</b>					
Race	16	...	2	82	5
Ethnicity	16	...	...	84	5
Age	47	...	8	45	5
Birth facility	63	2	5	31	8
City/town of residence	18	...	2	81	5
County of residence	16	...	2	82	5
State of residence	18	...	2	81	5
<b>Payer status</b>					
Medical insurance	72	...	7	21	6
Medicaid participant	54	...	17	29	8
WIC participant	27	2	27	44	8

*Note.* MMR=maternal medical record; IMR=infant medical record; PNR=prenatal record; WS=parental birth certificate worksheet; WIC=Women, Infants, and Children Program.

**TABLE 2—Electronic Birth Certificate Primary Data Sources for Medical History and Outcome Information: New Jersey Maternity Facilities, 1997–1998 (n=66)**

	MMR, %	IMR, %	PNR, %	WS, %	Missing, No.
<b>Medical history</b>					
Gravidity	35	...	52	14	9
Parity	33	2	51	14	10
Live-born offspring now living	32	...	54	14	10
Live-born offspring not living	31	...	53	16	9
Date of last live birth	31	...	57	12	9
Month prenatal care began	22	...	64	14	9
Location of prenatal care	22	...	66	12	9
Most recent menstrual period date	22	...	66	12	9
No. of prenatal visits	24	...	62	14	9
Risk factors	32	...	56	12	10
Obstetric procedures	48	...	41	10	9
Previous cesarean delivery	26	...	64	10	9
Tobacco use	30	...	56	15	6
Alcohol use	30	...	56	15	6
Drug use	30	...	56	15	6
Weight gain	31	...	56	13	6
<b>Maternal outcomes</b>					
Method of delivery	86	...	2	12	9
Maternal complications	90	2	...	9	10
<b>Infant outcomes</b>					
Delivery outcome	75	9	...	16	11
Birthweight	54	32	...	14	10
Sex	63	23	...	14	10
Apgar scores	65	21	...	14	10
Gestational age estimate	53	33	...	14	9
Abnormal conditions	25	61	...	14	10
Infant diagnosis	12	70	...	18	10
Plurality	70	14	...	16	10
Birth order	68	14	...	18	10
Transfer from other facility	14	72	...	14	9
Days in NICU	5	70	...	25	27
Discharge date	9	78	...	14	8
Length of stay	17	70	...	14	8
Feeding status	4	84	...	12	10
Future care provider	10	57	2	32	7
Final infant status	7	80	...	13	7

*Note.* MMR=maternal medical record; IMR=infant medical record; PNR=prenatal record; WS=parental birth certificate worksheet; NICU=neonatal intensive care unit.

## Discussion

There is currently a trend favoring the use of computerized data for both medical records and collection of health statistics. Advantages include timely data transmission, ability to ensure completeness of data entry via programmed electronic reminders to complete specific fields, and ability to rapidly analyze data for trends in health. Potential disadvantages include the ease of entering incorrect data and premature use of data for statistical analysis before accuracy can be verified.<sup>1-3</sup>

We identified 2 important potential sources of error in regard to vital statistics. The wide variety of primary data sources and data-entry personnel may introduce systematic bias into the information collected for an electronic birth certificate. However, any attempt to mandate potentially controversial standardized data collection and entry procedures may have adversely affected the original implementation of the electronic birth certificate in all maternity hospitals. The unavoidable consequence of this flexible approach to data procedures was that each facility developed its own system for data collection and entry.

It is not clear whether electronic birth certificate systems improve perinatal data accuracy. The use of different data sources is expected to have little impact on accuracy for fields such as identifiers and demographics. Data accuracy may be improved in other data fields.<sup>3</sup> Some studies have reported that use of check boxes or standardized electronic vital statistics data forms improves the accuracy of collected information.<sup>2,4,5</sup> However, these studies used the hospital medical record as the

“gold standard” for validating data. If the hospital record was not the original source of the data, validation efforts may have produced inaccurate results.

Wide disparities in data sources may partially moderate improvements in accuracy gained through direct input of information into the specific data fields of an electronic birth certificate.<sup>6</sup> There can be substantial discrepancies in vital statistics data such as gestational age, substance abuse, and prenatal care information, depending on the source of the data.<sup>1,2,4,6</sup>

The significant variations we have identified highlight the need to address standardization of data collection and data entry concurrently with the development of new electronic medical records. Because the temptation to use electronic data is so great, there is an accelerated need to develop standardized processes that should include validations of accuracy against primary data sources. Although potentially problematic for established electronic birth certificate programs, such a development may require extensive reengineering of perinatal data systems.<sup>3</sup> Additional research designed to determine optimal data sources is needed to assist existing and new electronic birth certificate programs. □

## Contributors

J. C. Smulian led the study design, implementation, and interpretation and the writing of the manuscript. All coauthors were involved in design of the study, selection of survey data fields, interpretation of results, and preparation of the manuscript. C. V. Ananth was also responsible for data analysis, and J. Don-

len was also responsible for implementation and administration of the survey.

## Acknowledgments

Dr Ananth was supported, in part, through grant 029553 from the Robert Wood Johnson Foundation awarded to the Center for Perinatal Health Initiatives. Additional support was provided through a grant from the American College of Obstetrics and Gynecology, District, Ill.

We gratefully acknowledge the support of the New Jersey Regional Maternal and Child Health Consortia, the participating maternity facilities, and the Health Research Education Trust of New Jersey, which made this project possible.

## References

1. Kirby RS. The quality of vital perinatal statistics data, with special reference to prenatal care. *Paediatr Perinat Epidemiol.* 1997;11:122-128.
2. Costakos DT, Love LA, Kirby RS. The computerized perinatal database: are the data reliable? *Am J Perinatol.* 1998;15:453-459.
3. Starr P, Starr S. Reinventing vital statistics: the impact of changes in information technology, welfare policy, and health care. *Public Health Rep.* 1995;110:534-544.
4. Gissler M, Teperi J, Hemminki E, Meriläinen J. Data quality after restructuring a national medical registry. *Scand J Soc Med.* 1995;1:75-80.
5. Piper JM, Mitchel EF, Snowden M, Hall C, Adams M, Taylor P. Validation of 1989 Tennessee birth certificates using maternal and newborn hospital records. *Am J Epidemiol.* 1993;137:758-768.
6. McDermott J, Drews C, Green D, Berg C. Evaluation of prenatal care information on birth certificates. *Paediatr Perinat Epidemiol.* 1997;11:105-121.